The EMBO Journal vol.5 no.10 pp.2577-2584, 1986

Purification of a factor specific for the upstream element of the adenovirus-2 major late promoter

V.Moncollin, N.G.Miyamoto[1], X.M.Zheng and J.M.Egly

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cédex, France

[1]Present address: Division of Biological Research, The Ontario Cancer Institute, Department of Medical Biophysics, University of Toronto 500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

Communicated by P.Chambon

Stimulation of *in vitro* transcription by the upstream element (UE) of the adenovirus-2 major late promoter (Ad2MLP) involves a specific *trans*-acting factor present in a HeLa whole-cell extract. By following its transcriptional stimulatory activity and its DNase I footprint on the Ad2MLP-UE, we have purified this factor to >10% purity and separated it from RNA polymerase B and the general transcription factors required for transcription from the Ad2MLP.

Key words: Ad-2MLP/RNA polymerase B/*in vitro* transcription/initiation factors

Introduction

The promoter region of genes transcribed by RNA polymerase B (II) are composed of several elements: (i) the cap site or start site element, which contains the bases coding for the 5' terminal nucleotides of the mRNA; (ii) the TATA box, a highly conserved A-T rich region which is located ~30 bp upstream from the start site and directs the transcription machinery to initiate at this site; (iii) the upstream elements which are located in the −40 to −110 region upstream from the start site and are required for efficient transcription and (iv) enhancers which stimulate transcription bidirectionally even when moved away from the promoter and located either upstream or downstream from the cap site (for references and reviews, see Breathnach and Chambon, 1981; Serfling *et al.*, 1985; Wasylyk, 1985).

In vitro and *in vivo* studies have shown that transcription factors interact with these promoter elements. At least one of these factors interacts specifically with the TATA box and start site region of different genes to form a stable pre-initiation complex (Davison *et al.*, 1983; Fire *et al.*, 1984; Parker and Topol, 1984a). This interaction occurs in the absence of RNA polymerase B and is stabilized by another factor (Egly *et al.*, 1984; Fire *et al.*, 1984) which appears to co-purify with actin [the 43 kd stimulatory transcription factor (STF) of Egly *et al.*, 1984 but see also Samuels and Sharp, 1986]. At least two other factors and RNA polymerase B are required to initiate specific RNA synthesis (Matsui *et al.*, 1980; Davison *et al.*, 1983; Fire *et al.*, 1984). Other sequence-specific factors which interact with upstream and enhancer elements have been identified and in some cases partially purified, e.g. those for the 21-bp repeat region of the SV40 early (Dynan and Tjian, 1983; Gidoni *et al.*, 1984, 1985; Barrera-Saldana *et al.*, 1985), the *Drosophila* heatshock hsp 70 gene (Parker and Topol, 1984b) and the SV40 enhancer (Wildeman *et al.*, 1984, 1985; Sassone-Corsi *et al.*, 1985).

Sequences located upstream from the TATA box (from −51 to −66/67) are required for efficient transcription from the adenovirus-2 major late promoter (Ad2MLP) both *in vivo* and *in vitro* (Hen *et al.*, 1982; Yu and Manley, 1984; Jove and Manley, 1984; Miyamoto *et al.*, 1984, 1985) and this upstream element (UE) is recognized specifically by a factor contained in HeLa cell extracts (Miyamoto *et al.*, 1984, 1985; Carthew *et al.*, 1985; Sawadogo and Roeder, 1985). To understand how this upstream element factor (UEF) stimulates initiation of transcription from the Ad2MLP, we have undertaken to reconstitute an *in vitro* transcription system from purified components. We report here the purification of the Ad2MLP upstream element factor (Ad2MLP-UEF). Concurrently, we have identified at least four general transcription factors (43 kd STF, BTF1, BTF2 and BTF3) which together with RNA polymerase B are necessary and sufficient for accurate initiation of transcription from either the wild-type or the UE-deleted Ad2MLP.

Results

The purification of the factor interacting with the upstream element of the Ad2MLP was followed by three methods: (i) selective stimulation of run-off transcription from a template which contains the Ad2MLP upstream element (Ad2MLP-UE) (Miyamoto *et al.*, 1984); (ii) DNase I footprinting (Galas and Schmitz, 1978) on the Ad2MLP-UE sequences (Miyamoto *et al.*, 1985); (iii) specific binding of Ad2MLP-UE DNA onto nitrocellulose (filter binding assay). The DNA template used in transcription assays was the double-promoter recombinant, pM34.97 (see Miyamoto *et al.*, 1984 and Figure 2A) which contains both an intact and an UE-deleted (between −34 and −97) Ad2MLP. Using *Acc*I-digested pM34.97, specific 309 and 212 nucleotide-long run-off transcripts are obtained for the deleted '34' and the intact '97' promoters, respectively (Figure 2A). Thus a selective stimulation of transcription from the '97' template can be used to follow the presence of the UEF during the purification. The purification steps are outlined in Figure 1 and Table I in which the amount of protein used for the footprinting, nitrocellulose binding or transcription assays is also indicated. Since the transcription reaction was not optimized at each step of purification for the amounts of the general transcriptional factors, it was not possible to define a unit of stimulatory activity for the upstream element factor.

Heparin−Ultrogel chromatography

The whole cell extract (WCE) was applied onto a Heparin−Ultrogel column, which yields two fractions absolutely required for accurate and efficient transcription from the Ad2MLP. The Heparin flow-through (HFT) contains a stimulatory transcription factor (43 kd STF), which has been previously

characterized (Egly *et al.*, 1984). Here we have mainly used the DE0.35 fraction [obtained by applying the HFT fraction on a DEAE-cellulose column (see Egly *et al.*, 1984)] as a source of 43 kd STF in the transcription assay. The 0.6 M KCl eluate from Heparin−Ultrogel (H0.6) contains RNA polymerase B and all the other general factors required for transcription from the cellular and viral promoters previously tested (SV40 early, Ad2MLP, Ad2E2, Ad5E3, rabbit β-globin, HSV thymidine kinase, chicken conalbumin; see Davison *et al.*, 1983; Miyamoto *et al.*, 1984 and data not shown).

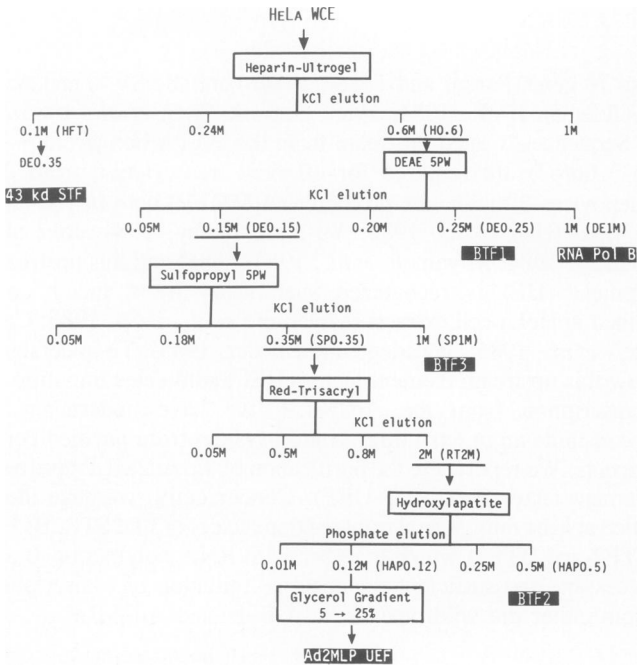

Fig. 1. Purification scheme of the Ad2MLP upstream element factor. The purification steps are described under Materials and methods. The abbreviations used are: kd, kilodaltons; STF, stimulatory transcription factor; BTF, RNA polymerase B transcription factor; UEF, upstream element factor.

The UEF appears to be present in the H0.6 fraction (Figure 2B, lane 1; compare the 309 nucleotide run-off from the '34' template and the 212 nucleotide run-off from the '97' template). Transcription from the '97' template was decreased to the level found with the '34' template in an *in vitro* competition assay in which H0.6 was incubated prior to addition of the pM34.97 plasmid with a DNA fragment (−260 to −34) containing the upstream element (see Miyamoto *et al.*, 1984). In addition increasing amounts of H0.6, but not of HFT, preferentially stimulated *in vitro* transcription from the '97' template (not shown).

The presence of the UEF in the H0.6 fraction was supported by DNase I footprinting (Figure 2C, lane 2) using the ^{32}P 5′ end-labelled coding strand from the wild-type Ad2MLP recombinant pM677 (see Materials and methods). A region of protection (P) was found between −67 and −50, as well as a major hypersensitive site (HS) at residue −69 and two additional hypersensitive sites at residues −73 and −71 (the footprint is described in detail in Miyamoto *et al.*, 1985). Although full protection was not achieved at this and the next step of the purification (most probably because of the presence of competing non-specific DNA binding proteins), the hypersensitive site at −69 constitutes a useful indicator of the specific binding of the UEF.

DEAE 5PW chromatography

The H0.6 fraction was separated into five fractions on a DEAE column by stepwise elution: the 0.05 M flow-through (DEFT), 0.15 M (DE0.15), 0.2 M (DE0.20), 0.25 M (DE0.25) and 1 M (DE1M) KCl fractions (Figure 1). When individual DEAE fractions were tested for *in vitro* transcriptional activity, no specific run-off RNA could be detected (Figure 2B, lanes 2 and 3 and

Table I. Purification summary of the Ad2MLP upstream factor

Purification step	Fraction	Volume (ml)	Protein concentration (mg/ml)	Total protein (mg)	Protein per assay (μg)			Total units[f]	
					Transcript.[b]	Footprint[c]	Retention[e]	Footprint	Retention
1	HeLa WCE[a]	420	7	2940	42	–	–	–	–
2	Heparin−Ultrogel (H0.6)	270	0.70	189	4.2	0.7[d]	–	–	–
3	DEAE-5PW (DE0.15)	40	0.45	18	1.8	0.9[d]	–	–	–
4	Sulfopropyl-5PW (SP0.35)	18	0.40	7.2	1.5	0.1	0.076	72 000	95 000
5	Red-Trisacryl (RT2M)	5	0.20	1	0.5	0.02	0.03	50 000	33 000
6	Hydroxylapatite (HAP0.12)	1.2	0.1	0.12	–	0.006	0.008	20 000	15 000
7	Glycerol gradient (pooled fractions)	2	0.001	0.002	–	0.003	0.003	700	700

[a]Prepared from 4.5 × 10^{10} HeLa cells (100 l of HeLa cells).
[b]Amount of protein which gives a level of *in vitro* transcription similar to the level obtained with the WCE.
[c]Amount of protein which yields 'full' footprint with 1 ng of DNA for each particular fraction.
[d]The footprint obtained with these fractions is not as strong as with the more purified fractions. Using higher amounts of proteins inhibits the footprint, probably due to binding of non-specific proteins. Hence these values cannot be directly compared with those given for the more purified fractions.
[e]Amount of protein required for the retention of 1 ng of DNA on the nitrocellulose filter.
[f]One unit = amount of protein required for full protection of 1 ng of DNA in the footprint assay or for retention of 1 ng of DNA in the filter binding assay.

data not shown). However, simultaneous addition of the DE0.15 and DE0.25 fractions to an incubation mixture containing calf thymus RNA polymerase B and 43 kd STF, yielded a specific RNA run-off (Figure 2B, lanes 4−7). No specific transcription was seen in a reaction containing the DE0.15 and DE0.25 fractions, but lacking either 43 kd STF or RNA polymerase B or both (Figure 2B, lanes 8−10). The DE1M fraction contains >97% of RNA polymerase B activity as determined by [^{3}H] α-amanitin binding (data not shown). Due to the instability of the enzyme in this fraction, we have used RNA polymerase B purified from calf thymus (Kédinger and Chambon, 1972).

The DE0.25 fraction was absolutely required to reconstitute transcription from either the '34' template or '97' template (compare Figure 2B, lanes 2 and 4−7). Using a competition assay with two different DNA templates (as described in Davison *et al.*, 1983), we have found that the DE0.25 fraction contains a factor (hereafter called BTF1) which binds to the TATA box region and is involved in the formation of a stable pre-initiation complex in the presence of the 43 kd STF (unpublished results). The DE0.15 fraction was also required to reconstitute *in vitro* transcription from both templates (compare Figure 2B, lane 3 and 4−7). Increasing the DE0.15 fraction concentration resulted in an increased transcription from the '97' template compared with the '34' template (see Figure 2B, lanes 4−7). Footprinting assays indicated that only the DE0.15 fraction contained the UEF (Figure 2C, lane 3 and results not shown). Since there was no transcription of the '34' template in the absence of the DE0.15 fraction (Figure 2B, lane 3), we conclude that this fraction contains at least one general transcription factor in addition to the UEF.

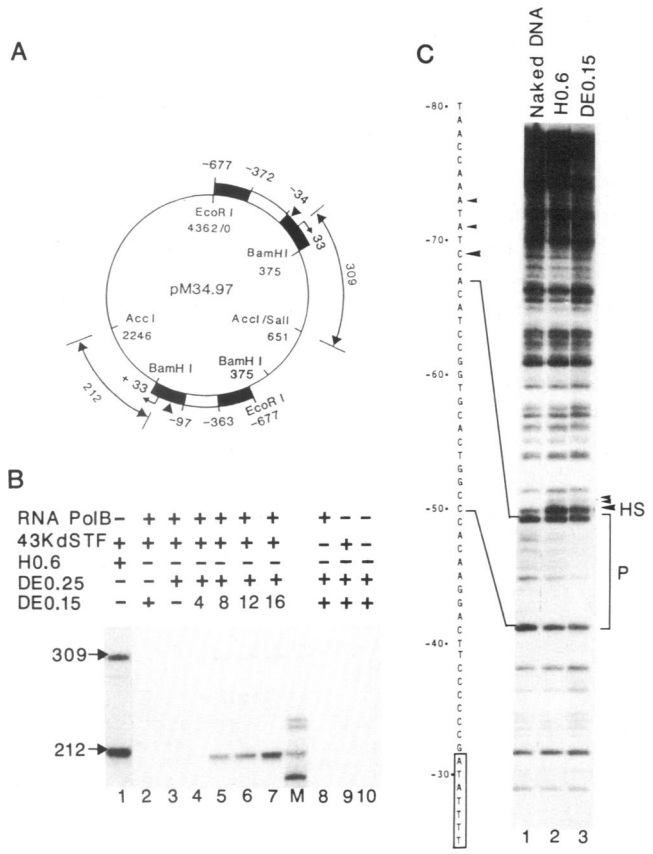

Fig. 2. (**A**) Structure of the Ad2MLP-containing recombinant used for *in vitro* transcription. The construction of the pM series has been described previously (Miyamoto *et al.*, 1984). The pM34.97 recombinant has two Ad2MLP regions (solid heavy lines) and was constructed by insertion of the repaired *Bam*HI Ad2MLP-containing fragment from pSVA97 (Hen *et al.*, 1982) into the *Pvu*II site of pM34. The '97' and '34' units contain deletion (blank space) from positions −97 to −363 and positions −34 to −372 of the Ad2MLP sequence respectively. (**B**) Reconstitution assays using the DEAE-fractions. Reactions were carried out as described under Materials and methods. Various combinations of the RNA polymerase B (0.2 μl) 43 kd STF (2 μl of DE0.35), H0.6 (6 μl), DE0.25 (4 μl) and DE0.15 (4 μl or as indicated) fractions were tested for specific transcription activity with 150 ng of the *Acc*I-digested pM34.97 plasmid as indicated at the top of the figure. The size of the '97' template (212 nucleotides) and '34 template (309 nucleotides) run-off are indicated by the arrows; M: size markers: ^{32}P end-labeled *Msp*I fragments of pBR322. (**C**) DNase I footprinting on the coding strand of the *Xho*I−*Bam*HI fragment of pM677 (PM677 has been previously described in Miyamoto *et al.*, 1984). The preparation of the footprinting template and the footprint assay are described under Matierals and methods. **Lane 1**, naked DNA; **lane 2**, 1 μl of H0.6; **lane 3**, 2 μl of DE0.15. P indicates the protected region. HS indicates the hypersensitive site.

Sulfopropyl 5PW chromatography

The DE0.15 fraction was applied to a sulfopropyl column and separated into four fractions by stepwise elution: the flow-through (SPFT), 0.18 M (SP0.18), 0.35 M (SP0.35) and 1 M KCl (SP1M) fractions (Figure 1). These fractions were first tested for the presence of the UEF using the DNase I footprinting assay. Only the SP0.35 fraction gave a 'full' footprint on the Ad2MLP-UE sequence (Figure 3A, lane 4). The upstream element was fully protected against DNase I digestion between −50 and −67 and the hypersensitive sites (HS) were very strong. (In some of the preparations the SP1M fraction gave a weak footprint as shown in lane 5 and in Figure 5A, lane 6.) The SP0.35 and the other sulfopropyl-eluted fractions were tested separately in a transcription assay (Figure 3B) containing the other essential transcriptional components; 43 kd STF, calf thymus RNA polymerase B and DE0.25 fraction. No specific RNA was ob-

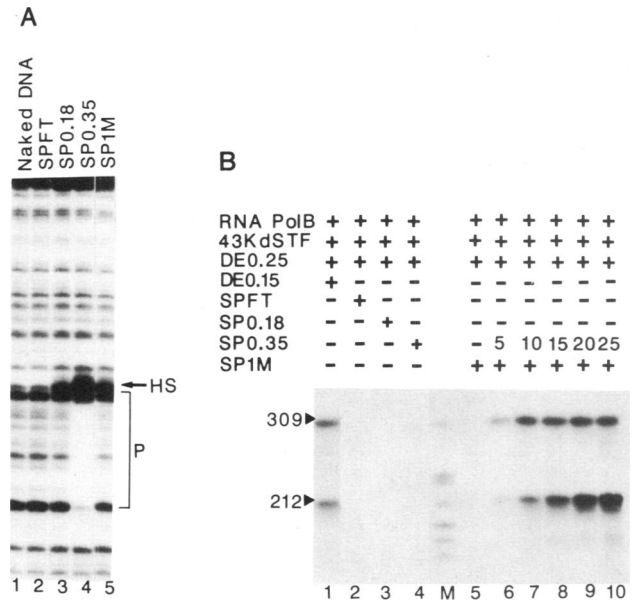

Fig. 3. Sulfopropyl chromatography of the DE0.15 fraction. (**A**) DNase I footprinting on the coding strand using SP fractions. **Lane 1**, naked DNA; **lane 2**, 5 μl of SPFT; **lane 3**, 5 μl of SP0.18; **lane 4**, 0.5 μl of SP0.35; **lane 5**, 5 μl of SP1M. (**B**) Reconstitution assays using the SP fractions. Reactions, DNA templates, size run-off and markers as in Figure 2B. Various combinations of the RNA polymerase B (0.2 μl) 43 kd STF (2 μl of DE0.35), DE0.25 (4 μl), DE0.15 (4 μl), SPFT (10 μl), SP0.18 (10 μl), SP0.35 (10 μl or as indicated) and SP1M (5 μl) fractions were tested for specific transcription activity as indicated at the top of the figure.

tained in the absence of the SP0.35 and SP1M fractions (Figure 3B, lanes 2–5). However, their presence restored accurate transcription of both '34' and '97' DNA templates (lanes 6–10). We conclude that SP0.35 and SP1M fractions contain each at least one general transcription factor (BTF2 and BTF3 respectively). Increasing the amount of SP0.35 fraction in the presence of a given amount of the SP1M fraction resulted in increasing transcription from the '97', but not from the UE-deleted '34' template (compare lanes 6–10). These results confirm that the Ad2MLP-UEF is present in the SP0.35 fraction.

Red Trisacryl chromatography

The SP0.35M fraction was chromatographed on a Red Trisacryl column (Figure 1). Most of the proteins present in the SP0.35M fraction were retained on this column, and were separated into four fractions by stepwise elution: flow through (RTFT), 0.5 M (RT0.5), 0.8 M (RT0.8) and 2 M KCl (RT2M) fractions. Footprinting analysis demonstrated that only the RT2M fraction contained the UEF (Figure 4A, lanes 5–8). These fractions were tested in a transcription assay containing RNA polymerase B and the three general transcription factors contained in the DE0.35 (43 kd STF), DE0.25 and SP1M fractions. Only the RT2M fraction restored transcription from the '97' as well as from the '34' template (Figure 4B, lane 4), indicating that the RT2M fraction contained some general transcription factor activity(s) present in the SP0.35 fraction. Furthermore, increasing amounts of RT2M stimulated preferentially transcription from the '97' template (Figure 4B, lanes 5–7), indicating the presence of the UEF in this fraction.

Hydroxylapatite chromatography

The RT2M fraction was applied onto a hydroxylapatite column and separated into three fractions by stepwise elution with a phosphate buffer: 0.12 M (HAP0.12), 0.25 M (HAP0.25) and

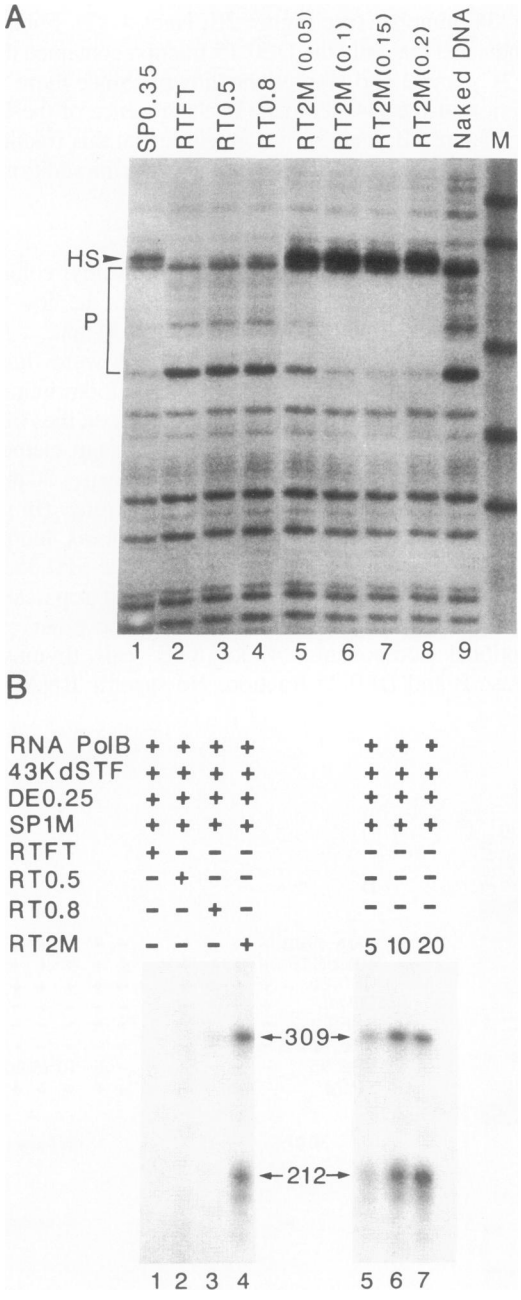

Fig. 4. Red-Trisacryl chromatography of the SP0.35 fraction. (**A**) DNase I footprinting on the coding strand using Red-Trisacryl fractions. **Lane 1**, 0.5 μl of SP0.35; **lane 2**, 15 μl of RTFT; **lane 3**, 15 μl of RT0.5; **lane 4**, 15 μl of RT0.8; **lanes 5–8** correspond to 0.05, 0.1, 0.15 and 0.2 μl of RT2M respectively; **lane 9**, naked DNA; **M**: size markers as in Figure 2B. (**B**) Reconstitution assays using Red-Trisacryl fractions. Reactions, DNA templates, size run-off as in Figure 2B. Various combinations of the RNA polymerase B (0.2 μl), 43 kd STF (2 μl of DE0.35), DE0.25 (4 μl), SP1M (5 μl), RTFT (15 μl), RT0.5 (15 μl), RT0.8 (15 μl) and RT2M (15 μl or as indicated) fractions were tested for specific transcription activity as indicated at the top of the figure.

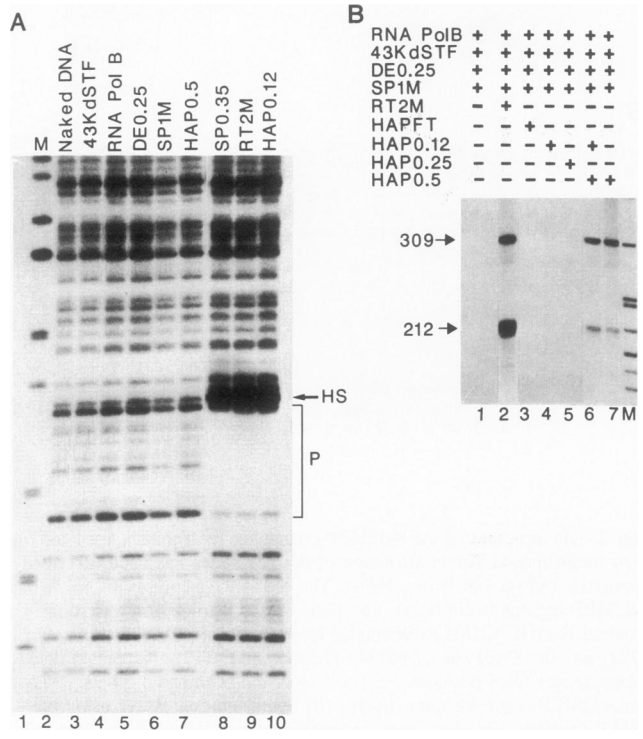

Fig. 5. (**A**) DNase I footprinting on the coding strand using the fractions necessary for accurate initiation of transcription. M: size markers as in Figure 2B; **lane 2**, naked DNA; **lane 3**, 1 μl of DE0.35; **lane 4**, 0.2 μl of calf thymus RNA polymerase B; **lane 5**, 4 μl of DE0.25; **lane 6**, 5 μl of SP1M; **lane 7**, 15 μl of HAP0.5; **lane 8**, 0.35 μl of SP0.35; **lane 9**, 0.1 μl of RT2M; **lane 10**, 0.1 μl of HAP0.12. (**B**) Reconstutition asays using the hydroxylapatite fractions. Reactions, DNA templates, size run-off and markers as in Figure 2B. Various combinations of the RNA polymerase B (0.2 μl) 43 kd STF (2 μl of DE0.35), DE0.25 (4 μl), SP1M (5 μl), RT2M (5 μl), HAPFT (20 μl), HAP0.12 (20 μl), HAP0.25 (20 μl) and HAP0.5 (20 μl) were tested for specific transcription activity as indicated at the top of the figure.

0.5 M (HAP0.5) fractions. Footprinting analysis showed that >80% of the UEF was eluted in the HAP0.12 fraction (Figure 5A, lane 10), whereas the rest was eluted in the HAP0.25 fraction (not shown). This fractionation resulted in a marked purification of the UEF factor. The HAP0.12 fraction contained only a few of the proteins loaded on the column (Figure 7, compare lanes 5 and 6), most of them being eluted in the HAP0.25 fraction (not shown).

The HAP0.5 fraction was essential to restore accurate transcription from both the '34' and '97' templates in the presence of RNA polymerase B, 43 kd STF, DE0.25 and SP1M fractions (Figure 5B, compare lane 7 for HAP0.5 with lanes 3, 4 and 5 for HAPFT, HAP0.12 and HAP0.25 fractions respectively). Thus the HAP0.5 fraction contains at least one general transcription factor (BTF2). Since the HAP0.12 fraction contains the UEF as revealed by DNase I footprinting it is clear that this factor can be separated from all general transcription factors and it is not necessary for accurate initiation of transcription from the Ad2MLP (see Figure 5B, lane 7). Addition of the HAP0.12 fraction to the transcription reaction resulted in only a moderate increase in the '97/34' template ratio (Figure 5B, lanes 6 and 7). By scanning the autoradiogram, we found that the '97/34' ratio was 1.9-fold higher after addition of HAP0.12 (lane 6 versus lane 7). As expected, none of the fractions necessary for accurate initiation of transcription from the '97' and the '34' templates protected the UE sequence from DNase I digestion (Figure 5A, lanes 3–7 for 43 kd STF, RNA polymerase B, DE0.25, SP1M and HAP0.5 fractions, respectively).

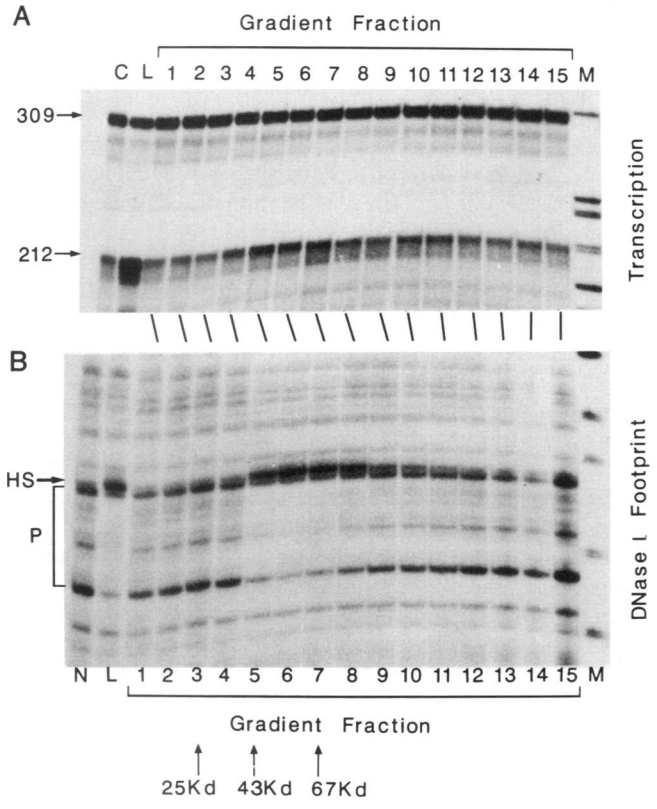

Fig. 6. Glycerol gradient analysis of SP0.35 fraction. 300 μl of the SP0.35 fraction was sedimented through a 5–25% glycerol gradient for 12 h at 60 000 r.p.m. in a SW60 rotor. The positions of chymotrypsinogen (25 kd), ovalbumin (43 kd) and serum albumin (67 kd) in parallel marker gradients are noted. (**A**) Gradient fractions were assayed after dialysis against buffer C for stimulation of transcription of the '97' template in the presence of a minimal amount (5 μl) of SP0.35 (see text) and 5 μl of SP1M, 4 μl of DE0.25, 2 μl of DE0.35 (43 kd STF) and 0.2 μl of RNA polymerase B. Reactions, DNA templates size run-off and markers as in Figure 2B. C (control): no other addition; **L**: addition of 15 μl of the SP0.35 fraction; **1–15**: addition of 25 μl of the gradient fractions. (**B**) DNase I footprinting on the coding strand using the gradient fractions. **N**: naked DNA; **L**: 0.5 μl of the SP0.35 fraction; **1–15**: 20 μl of the various gradient fractions.

Glycerol gradient centrifugation

The apparent mol. wt of the UEF was determined by glycerol density gradient centrifugation of the SP0.35 fraction. The glycerol gradient fractions were assayed for the UEF by both footprinting analysis and preferential stimulation of transcription from the '97' template. The basic transcription reaction contained the minimum amount of the SP0.35 fraction required to produce an equal basal transcription from the '97' and the '34' template (see Figure 3B, lanes 6 and 7), such that the addition of glycerol gradient fractions containing the UEF should result in a specific stimulation of transcription from the '97' template (Figure 3B, lanes 8–10). Both the stimulatory activity and the DNase I footprint of the UEF were found in the same glycerol gradient fractions (Figure 6A and B, lanes 5–7). This result strongly supports the conclusion that it is the UE binding factor which is responsible for the stimulation of transcription from the wild-type Ad2MLP. Centrifugation in parallel of marker proteins (Figure 6) indicated an apparent mol. wt of 55 ± 8 kd for the UEF.

Sedimentation on a glycerol gradient of the HAP0.12 fraction under the same conditions yielded an identical pattern of footprinting analysis (not shown; as in Figure 6B) and allowed us to eliminate most of the proteins (compare lanes 6 and 7, in Figure 7). The corresponding glycerol gradient fractions contained essentially two proteins migrating in the 54 and 57 ± 3 kd region (Figure 7, lane 7; the 65 kd band is an artefact due to silver staining procedure; for ref. see Tasheva and Dessev, 1983; Ochs, 1983).

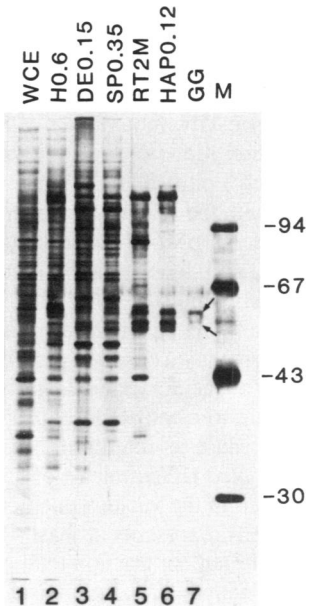

Fig. 7. SDS–polyacrylamide gel analysis of the various steps of purification. Aliquots (1 μg except for GG : 0.05) of the various protein fractions (as indicated) were electrophoresed and silver stained as described in Materials and methods. M: mol. wt markers. GG correspond to the pool of fractions 6 and 7 of the glycerol gradient ultracentrifugation of the HAP0.12 fraction. Small arrows indicate the position of the 54 and 57 kd bands which correlate with UEF binding activities.

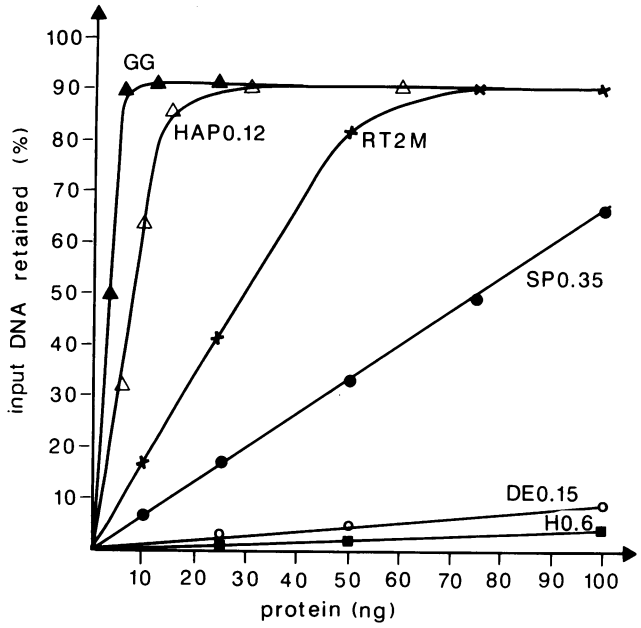

Fig. 8. Nitrocellulose filter binding assay. The binding of the UEF to the Ad2MLP upstream element was determined using the nitrocellulose filter assay described in Materials and methods (Hossenlopp *et al.*, 1974). Reactions were incubated for 15 min at 24°C with 2000 c.p.m. (~2 ng) of either pM677-X (*SacI*–*BamHI*) (wild-type upstream element) or pM1-X (*SacI*–*BamHI*) (mutated upstream element) DNA fragment and increasing concentrations of the various fractions. The difference, in c.p.m. retained, between pM677-X (*SacI*–*BamHI*) and pM1-X (*SacI*–*BamHI*) has been used to plot the curves. Approximately 10% of the mutated fragment is retained in the conditions where all the wild-type UE fragment is retained on the filter. GG corresponds to the pooled active fractions of the glycerol gradient step.

Estimation of the purity of Ad2MLP-UEF in the various fractions

We have used the specific retention of a DNA fragment containing the Ad2MLP-UE on nitrocellulose filters in an attempt to estimate the purity of the UEF in the various fractions (Figure 8). The 5′ end-labelled Ad2MLP *SacI*–*BamHI* fragment (−245/+33) of pM677-X (which has a wild-type UE, but a mutated TATA box, see Miyamoto *et al.*, 1984), was incubated with increasing concentrations of the various fractions and the solution filtered through a nitrocellulose filter. Non-specific retention was determined with the identical fragment of pM1-X which has the same sequence as pM677-X, but a double point mutation at −60 and −62, and does not exhibit any footprint on the Ad2MLP sequence (Miyamoto *et al.*, 1985). The difference (in c.p.m. retained on the filter) between retention of the pM677-X and pM1-X DNA fragment was used to plot the curves of Figure 8. The linear kinetics of the curves indicates that the binding of a single UEF molecule to a molecule of Ad2MLP-UE is most probably sufficient to cause its retention. Thus, using this assay and an assuming estimated UEF mol. wt of 55 kd, the minimum amount of UEF present in the various purification fractions can be calculated. The UEF represents at least 4% and 13% of the HAP0.12 and glycerol gradient fraction total proteins respectively. In good agreement with these values, we also calculated, from the amount of protein necessary to obtain a 'full' footprint on the UE, that the UEF represents 5.5% and 11% of the HAP0.12 and glycerol gradient fraction proteins, respectively (Table I).

Discussion

We have previously reported that the stimulation of *in vitro* transcription by the Ad2MLP-UE involves a specific factor present in a HeLa WCE (Miyamoto *et al.*, 1984, 1985). We report here the purification of this factor. In addition, the present purification procedure shows clearly that at least four general transcription factors, all distinct from the UEF, are required in addition to RNA polymerase B for accurate initiation of transcription from the Ad2-MLP, irrespective of the presence of the upstream element: the 43 kd STF which has been previously isolated (Egly *et al.*, 1984), a factor (BTF1) which is present in the DE0.25 fraction and binds to the TATA box region (T.Tamura, J.M.Egly and P.Chambon, unpublished results), and two other general transcription factors (BTF2 and BTF3) which are present in the HAP0.5 and SP1M fractions respectively (Figure 1). These results are in agreement with those of other groups, who have reported that several general transcription factors are required for initiation of the transcription from promoters of class B (II) genes (Matsui *et al.*, 1980; Tsai *et al.*, 1981; Samuels *et al.*, 1982; Davison *et al.*, 1983; Dynan and Tjian, 1983; Ackerman *et al.*, 1983; Parker and Topol, 1984a; Fire *et al.*, 1984).

Starting from a HeLa WCE containing 3 g of protein we have generally recovered ~2 μg of protein in the most purified fraction (glycerol gradient). The Ad2MLP-UEF displayed an apparent mol. wt of 55 ± 8 kd as judged by centrifugation in a glycerol gradient. Two major proteins present in the glycerol gradient fraction containing UEF migrated on an SDS polyacrylamide gel in the 55 kd region (Figure 7, lane 7). Using a nitrocellulose filter binding assay, we estimated the purity of the UEF to be 4% and 13% of the HAP0.12 and glycerol gradient fractions respectively. Using DNase I footprinting as an assay, similar values of 5.5 and 11% respectively were obtained. These values are likely to be minimal estimates. First an unknown fraction of the purified factor may have been inactivated (with respect to binding to DNA) during the purification, particularly in the last steps where the factor is present at very low concentration (see Table I). Secondly, the efficiency with which the UEF–DNA complex is retained on the filter is unlikely to be 100%. Thirdly, at the concentrations used in the present assay the binding of one molecule of UEF to the DNA may require the presence of more than one free molecule in solution. Thus the purity of the UEF in the glycerol gradient fraction may be >20% and the UEF corresponds most probably to one of the two proteins of 54 and 57 kd seen in lane 7 of Figure 7.

It is difficult to estimate by how many fold the UEF has been purified using the present purification scheme. However, from the amount of protein required to observe a full footprint on UE-DNA (Table I) or to give maximum retention on the nitrocellulose filter (Table I and Figure 8), the UEF appears to have been purified ~700-fold from the Heparin-Ultrogel to the glycerol gradient fraction. Since the transcription data suggest that an ~10-fold purification was achieved by Heparin chromatography (Table I), the overall purification is ~7000-fold.

A loss of transcription stimulatory activity was observed during the last steps of the purification of the UEF, although its specific binding activity was maintained (see Figure 5B, lanes 6 and 7, and Figure 5A, lane 10 respectively). This raises the possibility that an additional factor may have been lost during these steps. We were unable to restore the transcription activity by complementation with the other HAP fractions (data not shown). Further studies are in progress to determine whether this loss of stimulatory activity is due to dilution-related inactivation of the UEF and/or of other putative complementary factor(s), or to the concentration of an inhibitor. It is, however, clear that the binding of the UEF occurs in the absence of RNA

polymerase B and the other transcription factors (Figure 5A, lane 10).

The presence of a region of dyad symmetry in the Ad2MLP-UE has been described by us (Miyamoto *et al.*, 1985) and others (Sawadogo and Roeder, 1985; Carthew *et al.*, 1985). That the UEF possesses 2-fold rotational symmetry was strongly suggested by the pattern of dimethyl sulfate methylation protection experiments which have shown that most of the Ad2MLP-UE G residues which exhibit decreased or enhanced methylation in the presence of the UEF were symmetrically situated in the two halves of the dyad symmetry (Miyamoto *et al.*, 1985). Thus the Ad2MLP-UE may interact with either a protein dimer or a monomeric protein exhibiting an internal symmetry. The alternative possibility that each of the two halves of the UE independently binds one molecule of factor is unlikely as we have never observed independent protection of half of the dyad symmetry in footprint experiments carried out at limiting factor concentrations. At the present time we favor the hypothesis that the UEF may be a monomeric protein molecule possessing an internal symmetry because the glycerol gradient experiments shown in Figure 6 suggest the mol. wt of the putative factor to be the same as that seen by denaturing SDS polyacrylamide gel electrophoresis (Figure 7). Kinetic studies are in progress to investigate further whether the UEF is a monomeric or dimeric molecule.

Materials and methods

Purification of the upstream element factor of Ad2MLP

The purification scheme is shown in Figure 1. All procedures were carried out at 4°C. A HeLa WCE (70 ml) (Manley *et al.*, 1980) was prepared from 1.5×10^{10} HeLa cells and applied onto a column (4 × 10 cm) of Heparin–Ultrogel-A4R (IBF, Villeneuve-la-Garenne, France) (Davison *et al.*, 1983) equilibrated with buffer A [50 mM Tris-HCl pH 7.9, 100 mM KCl, 5 mM $MgCl_2$, 0.5 mM dithiothreitol (DTT) and 17.4% glycerol]. The column was washed (60 ml/h) sequentially with two column volumes of buffer A (HFT fraction), 1.5 column volumes of buffer A containing 0.24 M KCl, and finally 1.5 column volumes of buffer A containing 0.6 M KCl (H0.6 fraction). The HFT peak was then applied to a DEAE-cellulose (Whatman DE 22) column (2.5 × 8 cm) equilibrated in buffer A. After washing the column with two volumes of buffer A, the fraction containing the transcription stimulatory activity (43 kd STF, see Egly *et al.*, 1984) was eluted with buffer A containing 0.35 M KCl (DE0.35 fraction). This fraction was dialyzed against buffer B (10 mM Tris HCl pH 7.9, 50 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT, 17.4% glycerol), and used as source of STF in all transcription assays. The H0.6 fraction was dialyzed against buffer C (50 mM Tris HCl pH 7.9, 50 mM KCl, 0.5 mM DTT, 0.1 mM EDTA and 8.7% glycerol) and assayed for transcriptional activity in the presence of the 43 kd STF. The transcriptionally active H0.6 fractions obtained from three preparations were pooled and applied onto a preparative h.p.l.c. DEAE 5PW column (2.15 × 15 cm) (generous gift from Toyo-Soda, Japan) equilibrated in buffer C. After washing (2 ml/min) with two column volumes of buffer C (DEFT fraction) the absorbed proteins were successively eluted with 0.15 M KCl (DE0.15 fraction), 0.2 M KCl (DE0.2 fraction), 0.25 M KCl (DE0.25 fraction) and 1 M KCl (DE1M fraction) in buffer C. Each eluted fraction was dialyzed against buffer C and the DE0.15 fraction was then applied onto a h.p.l.c. sulfopropyl (SP, 5PW, TSK) column (7.5 × 75 mm) equlibrated in buffer C. The column was washed (0.6 ml/min) with two column volumes of buffer C (SPFT fraction) and proteins were eluted with 0.18 M KCl (SP0.18 fraction), 0.35 M KCl (SP0.35 fraction) and 1 M KCl (SP1M fraction) in buffer C. Each fraction was dialyzed against buffer C containing 20% glycerol and the SP1M fraction was stored in 0.2 ml aliquots at −80°C. The SP0.35 peak was then applied onto a column (2 ml volume in a Treff tip) of Red Trisacryl [Red 120 (Sigma) was immobilized on Trisacryl GF2000 (generous gift from IBF, France)] and the column was washed (8 ml/h) with five column volumes of buffer C (RTFT fraction), and proteins were eluted with five column volumes each of 0.5 M KCl (RT0.5 fraction), 0.8 M KCl (RT0.8 fraction) and 2 M KCl (RT2M fraction) in buffer C. Fractions of 0.5 ml were collected and the RT2M fraction was dialyzed against buffer D (10 mM potassium phosphate pH 7, 0.01 mM $CaCl_2$, 0.5 mM DTT, 8.7% glycerol); the RT2M fraction was applied onto a column (0.5 ml volume in a Treff tip) of hydroxylapatite (BDH, UK) equilibrated in buffer D. The column was washed (5 ml/h) with five column volumes of buffer D (HAPFT fraction) and proteins were eluted with eight column volumes each of 0.12 M (HAP0.12 fraction), 0.25 M (HAP0.25 fraction), 0.5 M (HAP0.5 fraction) potassium phosphate pH 7 in buffer D. Fractions of 0.3 ml were dialyzed and 0.3 ml of HAP0.12 fraction were sedimented through 5–25% glycerol gradients for 12 h at 60 000 r.p.m in a SW60 rotor. The fractions were stored at −80°C in 20% glycerol buffer.

In vitro transcription

The *in vitro* transcription run-off assay consisted of a 15 min pre-incubation at 24°C in a 16–40 μl reaction volume containing 50 mM Tris-HCl pH 7.9, 6 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM DTT, 50 mM KCl, 8% glycerol, 150 ng of the *Acc*I-digested pM34.97 plasmid and protein fractions as described in figure legends. Each transcription assay contains the 43 kd STF [2 μl (2 μg) of DE0.35 fraction; Egly *et al.*, 1984] and 0.2 μl (0.002 units) of partially purified calf thymus RNA polymerase B except otherwise indicated. After the pre-incubation period, 2 μl of a mixture containing 6 mM $MgCl_2$, 50 mM KCl, 2.5 mM each of ATP, GTP and UTP and 12.5 μM CTP containing 2–5 μCi [α-^{32}P]CTP was added to start RNA synthesis. Reactions were stopped after 45 min at 24°C, and the RNA run-off transcripts were purified and analyzed on 5% acrylamide 8.3 M urea gels as described in Davison *et al.* (1983).

DNAse I footprinting

DNase I footprinting assay was performed as previously described (Miyamoto *et al.*, 1985) and consisted of a 10-min pre-incubation at 24°C in a 18 μl reaction volume containing variable amounts of the protein fraction, ~1 ng (10 000 c.p.m.) of the *Xho*I–*Bam*HI fragment of pM677 (5′ end labelled on the RNA coding strand with [γ-^{32}P]ATP and T4 polynucleotide kinase; Maniatis *et al.*, 1982), 30–50 ng of unlabelled carrier pBR322 DNA fragments, 30 mM Tris HCl pH 7.9, 6 mM $MgCl_2$, 50 mM KCl, 0.07 mM EDTA, 0.3 mM DTT and 7% glycerol. After the pre-incubation 2 μl of buffer C containing 20 ng of DNase I (Worthington) was added and the reaction mixtures were incubated at 24°C for 2 min. Reactions were stopped and DNA digestion products were purified by phenol (0.05 M Tris-HCl, pH 7.9-saturated) extraction followed by ethanol precipitation and analyzed on 8% acrylamide 8.3 M urea gels followed by autoradiography.

Filter binding assay

The filter binding assay was done as described by Hossenlopp *et al.* (1974). The various protein fractions were incubated with 2 ng (2000 c.p.m.) of either pM677-X (*Sac*I–*Bam*HI) or pM1-X (*Sac*I–*Bam*HI) DNA fragment at 24°C for 15 min in buffer C containing 100 μg/ml bovine serum albumin and 10 μg/ml poly(dI.dC) for H0.6, DE0.15 and SP0.35 fractions or 2 μg/ml poly(dI.dC) for RT2M and HAP0.12 fractions in a final volume of 22 μl. The solution was then diluted to 200 μl with buffer C containing 100 μg/ml bovine serum albumin, filtered through a nitrocellulose filter (Millipore HAWP, 25 mm diameter) with gentle suction. The filter was washed with 4 × 1 ml of the same buffer, dried and the radioactivity determined in a liquid scintillation counter. Blanks without proteins were treated simultaneously and the blank values (2% of the input DNA fragment) were subtracted.

Other methods

Protein concentration was determined by a Coomassie blue staining method (McKnight, 1977). Protein gel electrophoresis was on 9% SDS-polyacrylamide gel and staining was carried out by either the Coomassie blue or the silver staining (Wray *et al.*, 1981) method.

Acknowledgements

We are indebted to P.Chambon for useful suggestions and discussions. We thank J.L.Plassat and A.Kempf for excellent technical assistance, M.Acker and J.L.Weickert for HeLa cell cultures, B.Boulay for photography, C.Werlé for illustrations, B.Wasylyk and I.Davidson for critical reading of the manuscript and E.Boschetti (IBF, France) and Y.Kato (Toyosoda, Japan) for synthesizing and providing chromatographic supports. N.G.M. is the recipient of a Post-Doctoral Fellowship from the Medical Research Council of Canada. X.M.Z. is the recipient of the CESTA (Paris) Biotechnology programme. This work was supported by the Action Incitative de Valorisation du CNRS (85-86), the Action MIR-Société IBF (85-86), and the Fondation pour la Recherche Médicale.

References

Ackerman,S., Bunick,D., Zandomeni,R. and Weinmann,R. (1983) *Nucleic Acids Res.*, **11**, 6041–6064.

Barrera-Saldana,H.A., Takahashi,K., Vigneron,M., Wildeman,A.G., Davidson,I. and Chambon,P. (1985) *EMBO J.*, **4**, 3839–3849.

Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.

Carthew,R.W., Chodosh,L.A. and Sharp,C.A. (1985) *Cell*, **43**, 439–448.

Davison,B.L., Egly,J.M., Mulvihill,E.R. and Chambon,P. (1983) *Nature*, **301**, 680–686.

Dynan,W.S. and Tjian,R. (1983) *Cell*, **35**, 79–87.

Egly,J.M,. Miyamoto,N.G., Moncollin,V. and Chambon,P. (1984) *EMBO J.*, **3**, 2363–2371.
Fire,A., Samuels,M. and Sharp,P.A. (1984) *J. Biol. Chem.*, **259**, 2509–2516.
Galas,D. and Schmitz,A. (1978) *Nucleic Acids Res.*, **5**, 3157–3170.
Gidoni,D., Dynan,W.S. and Tjian,R. (1984) *Nature*, **312**, 409–413.
Gidoni,D., Kadonaga,J.T., Barrera-Saldana,H., Takahashi,K., Chambon,P. and Tjian,R. (1985) *Science*, **230**, 511–517.
Hen,R., Sassone-Corsi,P., Corden,J., Gaub,M.P. and Chambon,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7132–7136.
Hossenlopp,P., Oudet,P. and Chambon,P. (1974) *Eur. J. Biochem.*, **41**, 397–411.
Jove,R. and Manley,J.L. (1984) *J. Biol. Chem.*, **259**, 8513–8521.
Kédinger,C. and Chambon,P. (1972) *Eur. J. Biochem.*, **28**, 283–290.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, New York.
Manley,J.L., Fire,A., Cano,A., Sharp,P.A. and Gefter,M.L. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3855–3859.
Matsui,T., Segall,J., Weil,P.A. and Roeder,R.G. (1980) *J. Biol. Chem.*, **255**, 11992–11996.
McKnight,G.S. (1977) *Anal. Biochem.*, **78**, 86–92.
Miyamoto,N.G., Moncollin,V., Wintzerith,M., Hen,R., Egly,J.M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 8779–8799.
Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1985) *EMBO J.*, **4**, 3563–3570.
Ochs,D. (1983) *Anal. Biochem.*, **135**, 470–474.
Parker,C.S. and Topol,J. (1984a) *Cell*, **36**, 357–369.
Parker,C.S. and Topol,J. (1984b) *Cell*, **37**, 273–283.
Samuels,M., Fire,A. and Sharp,P.A. (1982) *J. Biol. Chem.*, **257**, 14419–14427.
Samuels,M. and Sharp,P.A. (1986) *J. Biol. Chem.*, **261**, 2003–2013.
Sassone-Corsi,P., Wildeman,A.G. and Chambon,P. (1985) *Nature*, **313**, 458–463.
Sawadogo,M. and Roeder,R.G. (1985) *Cell*, **43**, 165–175.
Serfling,E., Jasin,M. and Schaffner,W. (1985) *Trends Genet.*, **1**, 224–230.
Tasheva,B. and Dessev,G. (1983) *Anal. Biochem.*, **129**, 98–102.
Tsai,S.Y., Tsai,M.J., Kops,L.E., Minghetti,P.P. and O'Malley,B.W. (1981) *J. Biol. Chem.*, **256**, 13055–13059.
Wasylyk,B. (1986) In Reznikoff,W. and Gold,L. (eds), *Maximizing Gene Expression*. Butterworths Publ. Group, London, pp. 79–99.
Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129–3133.
Wildeman,A.G., Zenke,M., Grundström,T., Schatz,C., Matthes,H., Wintzerith,M., Takahashi,T. and Chambon,P. (1985) *Mol. Cell. Biol.*, in press.
Wray,W., Boulikas,T., Wray,V.P. and Hancock,R. (1981) *Anal. Biochem.*, **118**, 197–203.
Yu,Y.T. and Manley,J.L. (1984) *Nucleic Acids Res.*, **12**, 9309–9321.

Received on 12 June 1986; revised on 14 July 1986